Strasbourg talk

Many of the functional attributes of retroviruses now appear to depend upon a domain of several hundred base pairs present at both ends of viral DNA. This domain, referred to hereafter as the long terminal repeat or L.T.R., is created during synthesis of retroviral DNA by the remarkable fusion of sequences derived from both the 5' and 3' ends of viral RNA. Since the discovery of LTRs about 3 years ago, much has been learned about their synthesis, structure, and multiple functions, particularly with recombinant DNA techniques. LTRs contain regulatory information crucial to the orderly progress of the virus life cycle. They can affect transcriptional activity of viral or heterologous DNA, apparently by multiple mechanisms, and these effects may be instrumental in oncogenesis.

My purpose here is to provide a brief review of these and other aspects of LTRs, saying more about what LTRs do than about how they are synthesized. I will begin with a schematic view of the life cycle which emphasizes the origin and role of the LTRs in the replicative mechanism. (First slide). The single stranded RNA subunit of all replication-competent retroviruses contains three coding domains; however, these genes and their products will not concern us further here. I direct your attention instead first to a short sequence, called R, represented as a solid box and present at both ends of viral RNA, and then to shaded and open boxes which represent those sequences - called U5 and U3 - unique to the 5' and 3' ends of viral RNA but repeated in DNA, by virtue of their inclusion in LTRs. Viral DNA is synthesized in a series of steps beyond the scope of today's discussion but the product is a linear duplex with LTRs composed of U5, R, and U3. The important feature to remember in thinking about the synthesis of an LTR is that the LTR domain is determined by priming sites. Thus, the first strand of DNA is primed by a host tRNA position at what will become the outer boundary of U5: and the second strand is initiated at a site which becomes the outer boundary of U3.

REPLICATION OF RETROVIRUSES

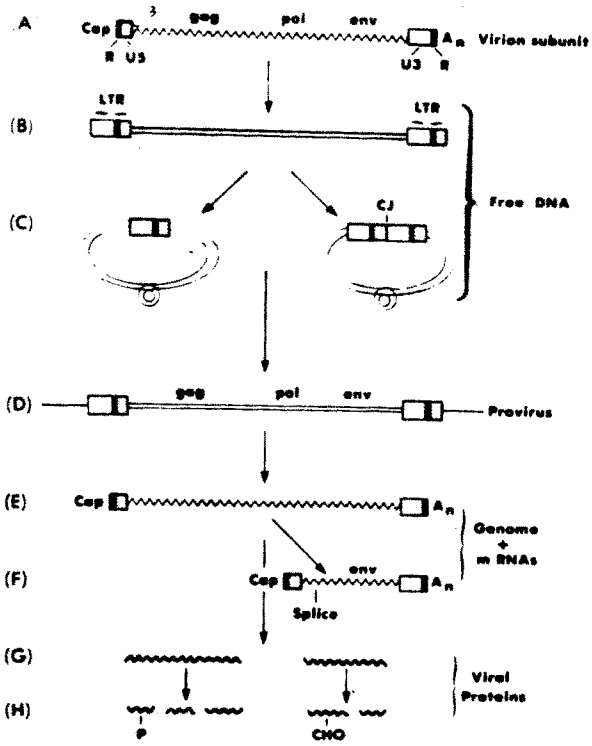


Fig 1

PROPERTIES OF LTRs

-FLANKED BY PRIMING SITES FOR VIRAL DNA.

-TERMINATE WITH INVERTED REPEATS.

-ENCODE INTEGRATION SITES 2 BP FROM EACH END.

-INCLUDE SITES AND SIGNALS FOR INITIATION AND POLYADENYLATION OF RNA.

Fig 2

STRUCTURE OF AN LTR

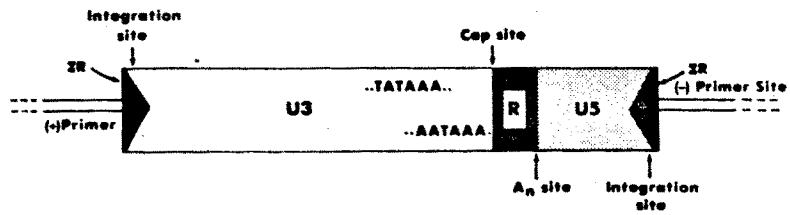


Fig 3.

Linear DNA can circularize to generate molecules with one or with two LTRs as shown here, or interesting aberrations of these structures. To produce an integrated provirus, one of these unintegrated forms - which one, is not known - is joined covalently to a host chromosome.

Three facts appear to be fairly firm about retroviral integration:

(1) there is little or no preference for integration sites in the host genome;

(2) integration sites in viral DNA are invariant, situated close to the ends of LTRs; and

(3) a short sequence of host DNA - 4, 5, or 6 bp, depending apparently upon the virus strain - is duplicated at the insertion site, forming a direct repeat which flanks the provirus.

The general structure of the provirus is appropriate for its function as a template for synthesis of nonpermeated viral RNA, which may in turn be spliced to form subgenomic mRNAs. It is also apparent from the diagram that LTRs are likely to figure prominently in the initiation, termination, or polyadenylation of transcripts.

Nucleotide sequencing of cloned LTRs from several retroviruses has confirmed some predicted features and revealed some unexpected ones. These are listed on the next slide (Fig. 2) and will be shown schematically on the one to follow (Fig. 3). LTRs are bounded on one side by the bonding site for host tRNA primer, and on the other by a polypurine tract that probably encodes the primer for the second DNA strand. LTRs end with short, inverted repeats; these are often imperfect and range from 5 to 20 or so bp's in length. Integration sites are invariably 2 bp's from each end, and the interior of the LTRs contains recognizable sites and sequences for initiation and polyadenylation of RNA conforming to the known structure of viral RNAs.

The next slide (Fig. 3) shows a more anatomical version of these findings. Notations across the top refers to features of relevant to the 5' LTR (i.e., the upstream or left hand LTR): the integration site; a Hogness Goldberg box about 30 bp from the probable initiation site for transcription; and the tRNA binding site. Insignia along the bottom refer to features relevant to the 3' (downstream) LTR: the polypurine primer site before the boundary; a canonical signal for poly A addition about 20 bp before the 3' end of R, and the integration site 2 bp from the end. (Although I have shown the polyadenylation signal here on the 5' side of the cap site - as it is in many viruses - some viruses encode this signal within R, creating a need to suppress this signal in the 5' LTR; how this is done is not known.)

At this point, I would like to consider some experimental evidence for the functional attributes of LTRs listed on the next slide (Fig. 4). In particular, I will discuss some recent work supporting the following contentions:

(1) LTRs can promote transcription, as deduced from experiments in which cloned LTRs have been joined to cloned heterologous genes in vitro, then reintroduced to cultured cells;

(2) LTRs have poorly defined properties which can enhance the efficiency of DNA transformation after microinfection of such recombinants;

(3) LTRs directly mediate certain regulatory events such as glucocorticoid control of mouse mammary tumor virus RNA synthesis:

(4) LTRs can affect the transcriptional activity of adjacent cellular genes, regardless of the arrangement of viral and host DNA, and such effects may contribute to oncogenesis, for example, in avian bursal lymphomas;

(5) the behavior of LTRs may be modified by their chromosomal context; and

(6) LTRs can act as sites for homologous recombination, facilitating the excision of proviral RNA from the chromosome.

Fig. 4

PROPERTIES OF LTRs

-CAN ACT AS PROMOTERS FOR LINKED GENES.

-ENHANCE EFFICIENCY OF DNA TRANSFORMATION AFTER MICROINJECTION.

-MEDIATE TRANSCRIPTIONAL REGULATION, e.g. BY GLUCOCORTICOID HORMONES.

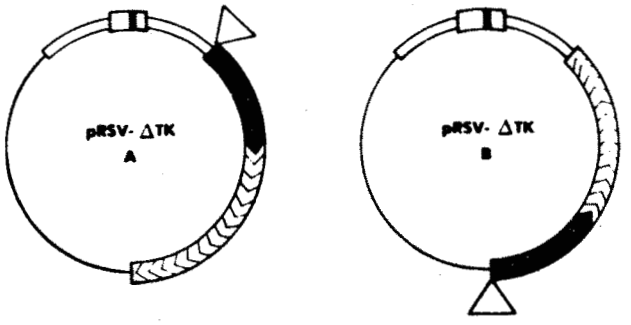
-AUGMENT TRANSCRIPTIONAL ACTIVITY OF ADJACENT CELLULAR GENES, e.g. IN BURSAL LYMPHOMAS.

-ACTIVITY INFLUENCED BY CONTEXT OF FLANKING DNA.

-MEDIATE PROVIRAL EXCISION BY HOMOLOGOUS RECOMBINATION.

Fig 5

AN LTR CAN PROMOTE TK EXPRESSION

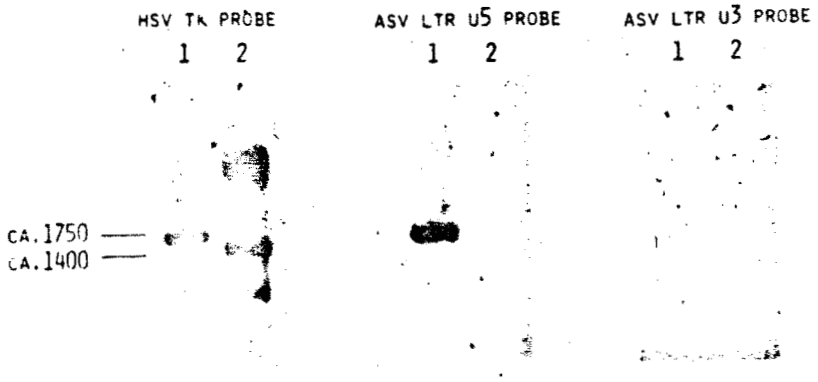


TK TRANSFORMATION AFTER MICROINJECTION

pRSV-ΔTK A	24%
pRSV-ΔTK B	0.2%
pΔTK	<0.1%
pTK	0.8%

Fig. 6

TK EXPRESSION PROMOTED BY ASV LTR



To test the promoter activity of the LTR of Rous sarcoma virus, Paul Luciw and Mario Capecchi microinjected the two plasmids illustrated on this slide (Fig. 5) into $\underline{tk}^-$ mouse L cells. Both plasmids contain a herpes $\underline{tk}$ gene, devoid of its own promoter, and a cloned fragment of retroviral DNA containing the LTR. In plasmid A, the LTR and the $\underline{tk}$ gene are in the same transcriptional orientation; in plasmid B, the tk fragment has been inverted. As tabulated at the bottom of the slide, a large proportion of cells receiving plasmid A were biochemically transformed to a $\underline{tk}^+$ phenotype, but few or none of those receiving plasmid B - or a plasmid with only the defective $\underline{tk}$ gene - were transformed. Interestingly, the frequency of transformation was significantly higher with plasmid A than with a plasmid ($\underline{ptk}$ containing an intact $\underline{tk}$ gene but no LTR. I shall return to this point in a moment.

To ask whether cells transformed by plasmid A were expressing the tk gene via the LTR, polyadenylated RNA from these cells was analyzed by gel electrophoresis and molecular hybridization. (Next slide; Fig. 6.) In each panel, RNA from cells transformed by plasmid A is present in lane 1. Probes for the $\underline{tk}$ gene and U5 - but not for U3 - detect a stable RNA species of the anticipated length from the transformed cells, a result consistent with initiation of synthesis within the LTR, near or at the end of U3. Lane 2 contains RNA from control cells transformed by a nondefective $\underline{tk}$ gene, without LTRs.

To investiate the surprisingly high efficiency of transformation by plasmids containing LTRs, Paul and Mario next injected cells with new plasmids (next slide; Fig. 7) containing the same fragment of Rous sarcoma virus DNA but a fully competent herpes $\underline{tk}$ gene in either orientation. Efficient transformation to a $\underline{tk}^+$ phenotype was now obtained using either plasmid; there was a greater than 20-fold increase over the frequency observed using a plasmid containing only a competent $\underline{tk}$ gene and no LTR. This experiment has also been

AN LTR INDIRECTLY ENHANCES TK TRANSFORMATION

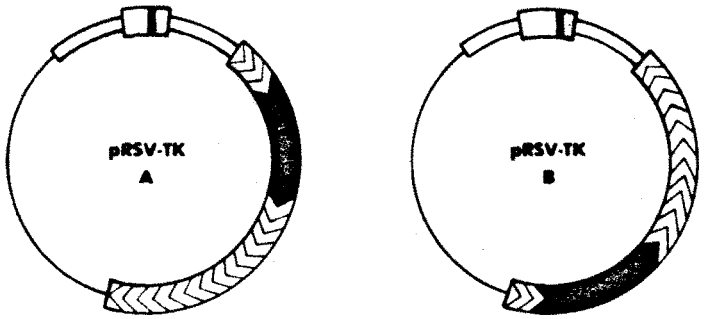
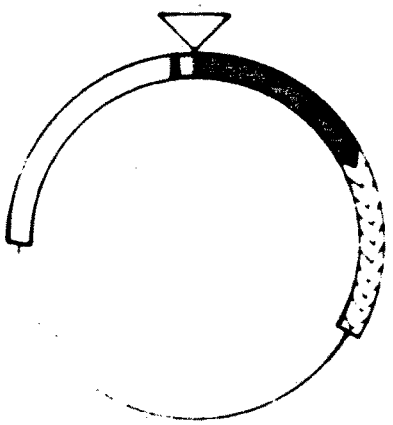


Fig 7

TK TRANSFORMATION AFTER MICRO INJECTION

pTK	0.8%
pRSV-TK A	23%
pRSV-TK B	19%

TK EXPRESSION FROM AN MMTV LTR RESPONDS TO GLUCOCORTICOIDS



\- +

MTV-TK

Fig 8

POSITIONS OF ALV PROVIRUSES WHICH ENHANCE EXPRESSION OF **C-MYC**

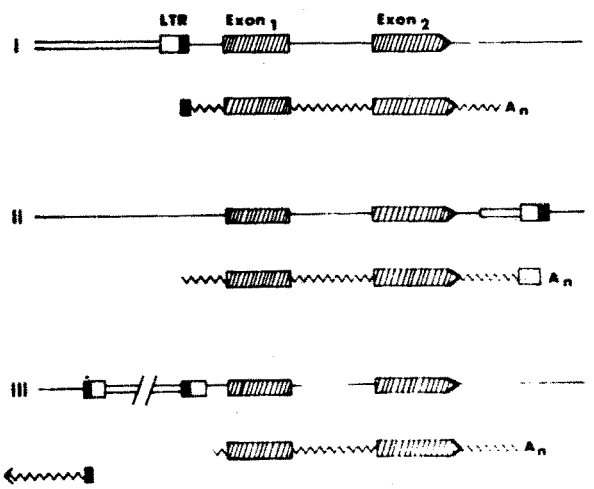


Fig. 9

performed with plasmids bearing an RSV DNA fragment containing only LTR sequences, with virtually identical results.

We do not know the mechanism of this enhancement. It could operate upon replication of free plasmids, upon integration, or upon transcription. If the effect is upon integration, it does not involve simply the donation of a good integration site in the LTR, since these plasmids do not preferentially integrate at the site used during natural viral infection. Similarly, if the effect is upon transcription, it is more mysterious than the simple provision of a superior promoter, since the effect is observed with the tk gene in both orientations. To distinguish these two types of transcriptional mechanisms I will refer - not entirely for humorous effect - to the provision of a good promoter as the force of an LTR and to more elusive enhancing effects as the charm of an LTR.

Expression of the mouse mammary tumor virus (MMTV) genome has been known for many years to be regulated at the transcriptional level by glucocorticoid hormones. To ask whether the force of MMTV LTR is hormonally modulated by interactions with the LTR itself (next slide; Fig. 8), John Majors has constructed plasmids with most of an LTR, cloned from a steroidally-responsive MMTV provirus, linked to a herpes tk gene without its promoter. Analysis of RNA from L cells into which this DNA was introduced [as a calcium phosphate precipitate] shows a dramatic increase in the amount of tk RNA after addition of hormone. Based on its size, this RNA is likely (but not yet proven) to be initiated within the MMTV LTR. Similar experiments performed with MMTV LTRs in other laboratories more clearly localize the site of initiation within the LTR. All agree that some component of the LTR is likely to mediate the hormone response.

We have been discussing the effects of LTRs upon expression of heterologous genes in unnatural settings, after construction of recombinants in vitro. But

the most dramatic evidence for such effects has emerged from recent studies of induction of B cell lymphomas after infection of birds by avian leukosis virus (or ALV). Viruses of this type, unlike agents such as RSV, produce tumors slowly, do not transform cultured cells, and lack a viral oncogene. Although virus-induced tumors contain ALV DNA, the DNA is often deranged so that coding domains are absent or unexpressed. However, viral DNA in each tumor is generally found in a common region of the host genome - a region identified by Hayward, Astrin and their colleagues as c-myc, the cellular homologue and predecessor to the oncogene of the transforming retrovirus, MC29,. Moreover, the level of c-myc RNA is markedly elevated in such tumors. What is responsible for the enhanced expression, "force" or "charm" or something else? (Next slide; Fig. 9.) In all of the tumors reported by Hayward et al - and in over half of those studied in our lab by Greg Payne - an ALV LTR is positioned on the 5' side (or upstream) c-myc - shown here as an interrupted cellular gene with two introns - in the same transcriptional orientation. In this arrangement, as shown at the top of the slide, the LTR can be predicted to act as an efficient promoter for c-myc. And, in fact, Hayward's group - and later, ours - found U5 (but not U3) sequences linked to c-myc sequences, in stable transcripts that are probably the spliced products of the putative primary transcripts diagrammed here.

However, in other tumors examined by Greg, "charm" seems to play a more significant role than "force". In one tumor, as diagrammed in the middle panel, a truncated provirus, containing little more than an LTR, is positioned on the 3' side of c-myc, in the same transcriptional orientation. The abundant c-myc transcripts in this tumor contain U3, but not U5 sequences; we do not know whether the normal c-myc promoter is employed to overproduce c-myc RNA in this case. In several other tumors (bottom panel) Greg has found an ALV provirus on the 5' side of c-myc but in the opposite transcriptional orientation. Again,

expression of c-myc is enhanced, but c-myc transcripts contain neither U3 nor U5. In at least one of those cases, an LTR promotes transcription away from c-myc into flanking cellular DNA.

How general are such phenomenon? Do they have equivalents in human cancers? Or in tumors induced by other viruses lacking oncogenes? Roel Nusse has recently used the strategy outlined in the next slide (Fig. 10) to assess the relevance of the ALV findings to carcinogenesis by MMTV. First he used viral probes to analyze Eco RI restriction digests of tumor DNA, to identify the rare tumor (T_1 with only a single provirus in addition to the several endogenous proviruses in normal mouse DNA (N). The right hand host-viral junction fragment was then cloned, a probe containing only single copy cellular sequences was prepared, and many tumors were analyzed for novel fragments - such as these - indicative of insertions or rearrangements of this such as that in T_2 - indicative of insertions or rearrangements of this unidentified region of the host genome. Thus far, at least 20% of mammary tumors are affected, but we have yet to define a transcriptionally active domain in this region or to identify unusual transcripts containing LTR sequences.

I would like to make two final brief points about LTRs, derived from yet another experimental context exemplifying insertion mutagenesis by retroviruses. (Next slide; Fig. 11.) In these experiments, a rat cell line (called B31) - transformed by a single RSV provirus - was morphologically reverted at low frequency after superinfection by murine leukemia virus, a non transforming retrovirus, when the latter's provirus was inserted between the promoter and the coding domain or between the splice sites (11) for the RSV transforming gene, src. In the two studied examples, both the RSV and MLV proviruses are in the same transcriptional orientation and the 5' RSV LTR remains transcriptionally active. Yet the first (5') MLV LTR does act to terminate or polyadenylate the RSV-

SEARCHING FOR INTEGRATION SITES INVOLVED
IN CARCINOGENESIS BY **MMTV**

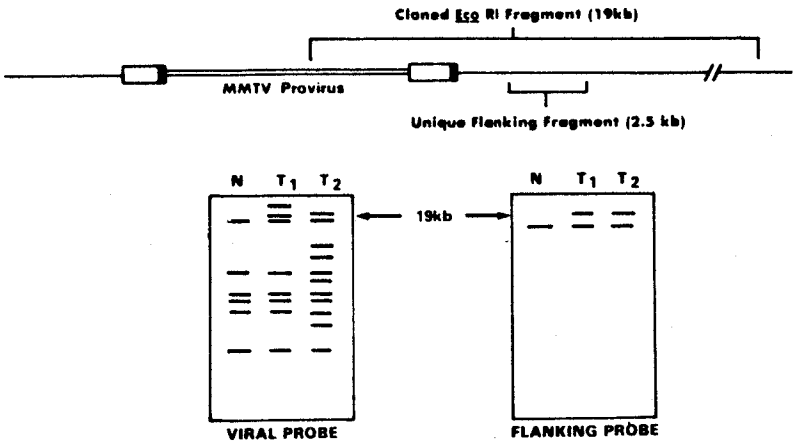


Fig 10

INSERTION MUTATION AND EXCISION

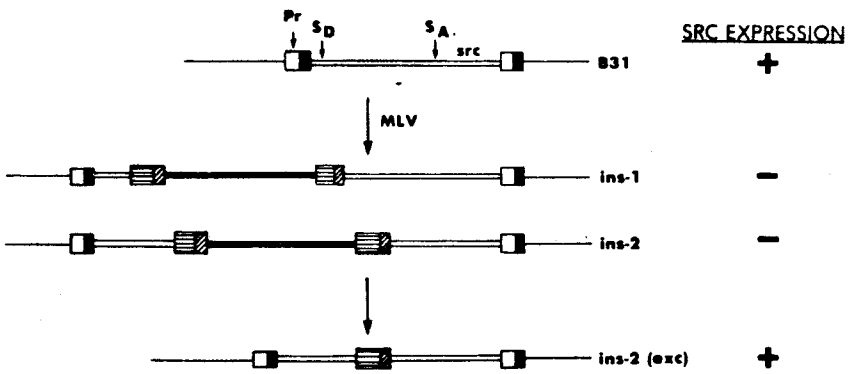


Fig 11

SUMMARY

1. LTRs CONTAIN SEQUENCES WHICH DETERMINE IMPORTANT STEPS IN THE VIRUS LIFE CYCLE.

2. LTRs MAY INFLUENCE TRANSCRIPTION BY MULTIPLE MECHANISMS ("FORCE" VS "CHARM")

3. LTRs MAY PROMOTE ONCOGENESIS BY ACTIVATING HOST GENES.

4. LTRs MAY SERVE AS SITES FOR HOMOLOGOUS RECOMBINATION.

Fig 12

initiated transcripts, the 3' MLV LTR does not promote expression of src, and no stable transcripts of src are found. In the second of these two mutants, retransformation occurs spontaneously at low frequency, with reappearance of normal src mRNA and protein. Most of the MLV provirus is now gone, but a single MLV LTR remains, suggesting that homologous recombination between LTRs has occured. The residual MLV LTR behaves like a neutral sequence in the src "intron", without apparent influence over transcription or processing. These results indicate two things (1) that the functions of LTRs may be influenced by their position (or context) in host chromosomes, and (2) that LTRs may play a phenotypically-important role as sites for homologous crossing over.

In summary then (last slide; Fig. 12) I wish to leave you with four points about LTRs:

(1) they contain sequences with important roles in lye cycle

(2) they influence transcription by multiple mechanisms

(3) they may promote neoplasia by activating host genes

(4) they may elininate proviruses by providing recombinational sites.